Word Learning in Children with Specific Language Impairment: Influence of Child and Word Characteristics

By

Mollee J. Pezold

Submitted to the graduate degree program in Speech-Language-Hearing: Sciences and Disorders and the Graduate Faculty of the University of Kansas in partial fulfillment of requirements for the degree of Master of Arts.

____
Chairperson Holly Storkel, Ph.D.

____
Nancy Brady, Ph.D.

____
Kristin Pedersen, M.A.

____
Mabel Rice, Ph.D.

Date Defended: March 30, 2017

The Thesis Committee for Mollee Pezold
Certifies that this is the approved version of the following thesis:

Word Learning in Children with Specific Language Impairment:
Influence of Child and Word Characteristics

____
Chairperson Holly Storkel, Ph.D.

Date approved: March 30, 2017

Abstract

Children with specific language impairment (SLI) have significantly poorer word learning ability than their same-age peers. It is not clear exactly where in the process of learning a new word these children struggle, but literature from experimental studies suggests that children's abilities in other areas and characteristics of the words themselves play a role. This study examined the influence of child characteristics (fast-mapping ability, phonological working memory, semantics, and language ability) and word characteristics (phonotactic probability, neighborhood density, and part of speech) on word learning outcomes in a clinical trial. Thirteen kindergarten children with SLI were taught vocabulary words through an interactive book reading treatment. Results showed that children who performed better on tasks of fast mapping and answering questions after listening to a short story generally learned more words following the treatment. Words from sparser phonological neighborhoods were generally learned by more children. Possible modifications to the treatment based on these findings are discussed.

Acknowledgements

I am grateful to my advisor, Holly Storkel, for her guidance, patience, and tireless editing. I would also like to thank my committee members, Nancy Brady, Kris Pedersen, and Mabel Rice, for their helpful feedback throughout this process.

Thank you to the children, families, and school districts that made this research possible through their participation. I would also like to acknowledge the financial support for this project, National Institutes of Health Grant DC012824, awarded to Holly Storkel.

For their assistance on this project and for their friendship, I thank the current and former members of the Word and Sound Learning Lab: Rouzana Komesidou, Breanna Krueger, Kelsey Flake, Veronica Fierro, Adrienne Latta, Simone Huls, Junko Maekawa, and Nicole Schuh.

Special thanks goes to Ali Sultani for his proofreading and moral support and to Sandie Pezold for a lifetime of encouragement.

Contents

Introduction

Specific language impairment (SLI) is a language disorder without a known cause that affects an estimated 7.4% of kindergarten children (Tomblin et al., 1997). Numerous studies have shown that children with SLI have difficulty learning new words. In a meta-analysis of 28 studies, Kan and Windsor (2010) found that children with SLI had significantly poorer word learning ability than age-matched typically developing peers, although this effect disappeared when they were compared to younger, language-matched peers. This suggests that children with SLI have the expected word learning performance given their language level, which is by definition delayed relative to their same-age peers. It is clear that children with SLI struggle in the process of learning new words, but it is not clear where in the word learning process these children do or do not struggle.

When a child first hears a new word, he or she must determine that the word is unknown and needs to be learned. This is likely accomplished based on the overall novelty of the word as well as an inability to call up the word in the child's mental lexicon. The child must then hold the word form in working memory while attempting to identify the referent or meaning of the word, either through visually searching for an unfamiliar referent in the environment or listening to the accompanying conversation to infer the meaning of the word form. This first step is called fast mapping. Characteristics of the words themselves may aid in children's fast mapping. The grammatical class of a word may aid children's learning at this stage, with more-concrete nouns learned more rapidly than less-concrete descriptors and verbs (e.g., Nelson, 1973). Other word characteristics of importance include phonotactic probability, which describes how likely a specific sound sequence is to occur, and neighborhood density, which describes the total number of words that differ from a word by only one phoneme, or sound. Studies generally find that typically developing children have better learning for novel words with low phonotactic probability from sparse neighborhoods (e.g., Storkel, Armbruster, & Hogan, 2006), likely because these words are distinct from other words children already know, aiding in the creation of a new representation.

Once the novel word form and referent have been identified via fast mapping, a representation must be created in memory, a process sometimes called extended mapping. Children must relate the sound sequence to other existing representations to create a new representation and store a semantic representation of the word's meaning. Children who already have large vocabularies are more successful in learning new words (e.g., Gathercole, Hitch, Service, & Martin, 1997). Gathercole and colleagues posit two possible reasons for this finding: children who know more words may have more stored sound sequences to which they can relate new word forms, or they may have more stored word meanings to which they can relate new semantic information. Likewise, recent research has shown better retention for words from dense neighborhoods one week later (Storkel & Lee, 2011), possibly because these words were integrated with the many similar words the children already knew.

The process of learning a new word is complex and consists of multiple steps. Exactly where this process breaks down in children with SLI is not clear. Although the finding that children with SLI perform worse on word learning tasks than their same-age peers is generally consistent across studies, there is also considerable variation across children and tasks. The goal of this study is to explore how child and word characteristics influence word learning to pinpoint relative strengths and weaknesses within the same group of children. This paper addresses these questions by examining data from an ongoing clinical trial of an interactive book reading vocabulary treatment for kindergarten children with SLI. The existing literature on SLI and results from the previously conducted study in the clinical trial will be used to establish the variables of interest for analyzing the data in the ongoing follow-up study of the clinical trial, which includes 13 kindergarten children with SLI. Throughout this paper the previously conducted and published trial is referred to as Study 1, and the ongoing follow-up clinical trial whose data are reported here is referred to as Study 2. Study 1 had a smaller sample size and divided children across different treatment groups; by analyzing data from Study 2, we will have more power to explore variables of interest.

Child Characteristics

Poorer word learning ability in children with SLI may be attributable to their impaired fast-mapping ability. Rice, Buhr, and Nemeth (1990) found that children with SLI had lower fast-mapping performance than both language-matched and age-matched peers. In another study comparing the fast-mapping performance of children with SLI to language- and age-matched peers, Beverly and Estis (2003) found that children with SLI were less likely to demonstrate disambiguation (i.e., mapping a nonword to an unfamiliar object) using the assumption of mutual exclusivity, especially when nonsense words were phonetically similar to real words. Estis and Beverly (2015) replicated this finding in preschoolers with SLI, and also found that school-age children with SLI did disambiguate phonologically distinct words as often as their same-age peers, but not phonologically similar words. These findings suggest that children with SLI may struggle with the fundamental step of mapping a new word to its referent, but may receive assistance from distinct phonological information (but see Gray, 2006).

Study 1 (the previously conducted interactive book reading treatment study; Storkel et al., 2017) did not include specific measures of fast mapping. However, all children completed the Semantic Subtest of the Diagnostic Evaluation of Language Variation (DELV; Seymour, Roeper, de Villers, & de Villers, 2005), which does include a small subset of items evaluating fast mapping of real and novel verbs. The overall DELV Semantic score was significantly correlated with the number of study words correctly defined after treatment (i.e., higher DELV Semantic scores predicted more words learned). We investigated the fast-mapping items separately in the Study 2 data to examine a possible relationship between fast mapping ability and word learning performance.

Phonological memory describes the ability to temporarily store strings of sounds. Researchers frequently measure phonological memory with tasks of nonword repetition and digit span (i.e., recalling a string of numbers). McKean, Letts, and Howard (2013) found that children with SLI had poorer nonword repetition than their peers and that nonword repetition ability predicted vocabulary in both groups, a finding replicated by Torrens and Yagüe (2016). Other studies have linked nonword repetition performance to fast-mapping performance in children with SLI (Alt & Plante, 2006; Jackson, Leitao, & Claessen, 2015). However, other studies have failed to find a relationship between phonological working memory and novel word learning (Gray, 2006; Hansson, Forsberg, Löfqvist, Mäki-Torkko, & Sahlén 2004). Regarding this discrepancy, Leonard (2014) noted that phonological memory (specifically nonword repetition ability) may be a related but not a causal factor in children's word learning performance. Leonard speculated that children who gain larger vocabularies may improve their phonological representations and therefore their nonword repetition performance. This might explain the fact that overall vocabulary is consistently found to be related to nonword repetition while initial word learning is not.

Regarding nonword repetition, Study 1 (Storkel et al., 2017) found that nonword repetition scores (measured by the Comprehensive Test of Phonological Processing, [CTOPP];

Wagner, Torgesen, Rashotte, & Pearson, 2013) significantly predicted the number of study words correctly defined following treatment. That is, children who scored higher in nonword repetition learned more words following the treatment. This discrepancy with previous research may be explained by differences in teaching paradigms; fast mapping studies typically teach nouns via a single picture and very simple sentences, whereas the current study taught words in context via interactive book reading. The interactive book reading context may place more demands on children's working memory.

As previously noted, overall vocabulary can index a variety of skills needed for word learning. In addition, overall vocabulary may influence the integration of new words with existing words. There are a number of standardized vocabulary tests that are a general index of vocabulary, but the relationship between these tests and children's word learning has been inconsistent. For example, Gray (2004) and Jackson, Leitao, and Claessen (2015) both found that receptive vocabulary (measured by the Peabody Picture Vocabulary Test [PPVT]) predicted performance on a fast-mapping task in children with SLI. However, other studies using the PPVT (Gray, 2003, 2006; Rice et al.,1990; Rice, Oetting, Marquis, Bode, & Pae, 1994) and one using the Receptive One-Word Picture Vocabulary Test (ROWPVT; Kiernan & Gray, 1998) failed to find a significant relationship between fast mapping and vocabulary test scores. In addition, all of these studies looked at children's initial word learning, which may differ from children's longer-term retention of words over time. Measures investigating children's semantic knowledge of words (e.g., knowledge of attributes of or relationships between words) rather than the number of words they can identify may clarify this relationship. Overall, semantics may be a relative area of strength for children with SLI; Gray (2004) found that children with SLI had better semantic than phonological knowledge for newly learned words (but see McGregor et al. [2013] for contrasting findings in adults with language impairment).

Study 1 (Storkel et al., 2017) included two measures of semantics: the DELV Semantics Subtest and the Clinical Evaluation of Language Fundamentals (CELF) Word Classes Subtest (Semel, Wiig, & Secord, 2003). As described previously, the DELV Semantics Subtest score was significantly correlated to children's word learning performance. This subtest includes items testing verb and preposition contrasts, quantifiers, and fast mapping with real and novel verbs. The CELF Word Classes Subtest, which tests children's knowledge of word relationships, was not correlated with children's word learning performance. The nature of these tests makes it difficult to compare results with previous studies that used the PPVT or ROWPVT. Investigating the relationship of these two vocabulary measures to children's word learning performance in Study 2 may shed more light on the influence of children's semantic knowledge on their long-term word learning performance.

One might expect that a child's language abilities may influence word learning. Results in the literature are mixed. On the one hand, in their meta-analysis, Kan and Windsor (2010) found that receptive language ability predicted group differences in word learning ability when comparing children with SLI to age-matched peers. On the other hand, in a study on novel word learning, Gray (2003) found that both receptive and expressive language test scores (measured by the Oral and Written Language Scales [OWLS]) did not predict word learning ability for children with SLI. In terms of Study 1, none of the language test measures were significantly correlated with word learning outcomes (Storkel et al., 2017).

Word Characteristics

For phonotactic probability, several studies have found that, like typically developing children, children with SLI show an advantage for words with more likely sound sequences (Gray & Brinkley, 2011; Gray, Brinkley, & Svetina, 2012; McKean et al., 2013). In contrast, Gray, Pittman, and Weinhold (2014) found that while typically developing children showed an advantage for words with high phonotactic probability in referent identification tasks (i.e., identifying the object that matched a word) and low phonotactic probability in naming tasks (i.e., saying the name of the word when presented with the matching object), children with SLI showed no such advantage. Other studies have found poorer word learning performance by children with SLI on low probability than high probability words in fast-mapping tasks (Alt & Plante, 2006) and in production but not recognition tasks (Alt, 2011). If children with SLI show different effects of phonotactic probability on word learning than do typically developing children, it could be caused by differences in their storage and retrieval of phonological representations.

For neighborhood density, Gray and colleagues (2014) found that, while young typically developing children benefitted from sparse neighborhoods for both referent identification and naming, children with SLI benefitted from dense neighborhoods for referent identification only. Gray and colleagues suggest that children with SLI formed less-specific representations of the words' form, which led to more holistic processing of the words and an advantage for words with more neighbors.

Phonotactic probability and neighborhood density were not significant predictors of the words children learned in Study 1 (Storkel et al., 2017). However, not all children in Study 1 received the optimal number of exposures to the words, which resulted in very few words learned for some children and may have made the effect of differences in word characteristics difficult to detect. In Study 2 (the ongoing current study), all children received the number of exposures to each word that was identified as optimal (i.e., 36 exposures to each word). This should result in better performance across all children and more power to detect the influence of these word characteristics. In addition, the interactive book reading paradigm used in this clinical trial may result in different effects of phonotactic probability and neighborhood density than have previous studies, which typically provided fewer exposures over a shorter amount of time.

Children with SLI, like typically developing children, generally learn nouns (i.e., object labels) more readily than verbs (i.e., action labels; Eyer et al., 2002; Oetting, Rice, & Swank, 1995; Rice et al., 1990; Windfuhr, Faragher, & Conti-Ramsden, 2002). Adjectives (i.e., attribute labels) have received less attention, but there is some evidence that children with SLI also learn adjectives more readily than verbs (Rice et al., 1990). Although this noun advantage appears to be relatively robust, there are contradictory findings, but these may be attributable to the stimuli used (Horohov & Oetting, 2004; Rice et al., 1994). The authors noted that both of these studies used storybooks chosen for their depictions of character actions, which meant that the verbs played a more central role in the stories than they typically would. In line with previous studies, Study 1 (Storkel et al., 2017) found that children showed better learning for nouns than verbs and moderate learning for adjectives.

Previous research on word learning in children with SLI has focused primarily on experimental tasks of initial word learning. Study 1 (Storkel et al., 2017) provided an initial investigation into children's learning over more exposures and a longer timeframe. Study 2 takes advantage of the initial study's finding that children showed the best learning after 36 exposures to a word. The purpose of the current analysis is to examine the factors (child and word characteristics) that influence word learning performance of children with SLI to explain variation in treatment response. Investigating the variation children's word learning in Study 2, where all children received the optimal number exposures to the study words, may provide further insights into the factors that influence word learning in children with SLI over time. This is a small, exploratory investigation involving only 13 children. The goal is to provide a preliminary investigation of promising predictors that could be confirmed in larger studies.

Research Questions

1. Do child characteristics (specifically, fast-mapping ability, phonological working memory, semantics, and language ability) relate to the number of words learned during treatment?
2. Do word characteristics (specifically, phonotactic probability, neighborhood density, and part of speech) relate to the percentage of children that learned each word?

Method

Study 1, an earlier phase of the preliminary clinical trial, established that children with SLI showed optimal learning after 36 elaborated exposures to a word, with no increased benefit for exposures beyond 36 (Storkel et al., 2017). This was 3 times the number of exposures received by children in the study by Justice, Meier, and Walpole (2005) using the same materials and similar procedures. Study 2 (the current study) expanded on this finding to determine the optimal dose (i.e., the number of exposures to a word within a reading session) and dose frequency (i.e., the number of times the book is read) of treatment for word learning. Possible conditions included: maximize dose regimen (9 exposures to the target word in each of 4 reading sessions per book); balanced regimen (6 exposures to the target word in each of 6 readings per book; this regimen was used in the earlier study); and maximize dose frequency regimen (4 exposures to the target word in each of 9 readings per book). All children received two of the three possible treatments—the balanced regimen and either the maximize dose or maximize dose frequency regimen. Analyses for this study were conducted on children's first of the two treatments, as the second treatment had not been completed for all children. Analyses were conducted on children in the balanced regimen (6 exposures and 6 readings) and in the maximize dose frequency regimen (4 exposures and 9 readings) because preliminary results, described further in Appendix A, suggested that children in the maximize dose regimen (9 exposures and 4 readings) did not show adequate learning. The children included in the analyses showed a wide range in the number of words learned (M = 6.46, SD = 4.99, range = 0-16), while the children in the excluded condition all showed little learning (M = 1.50, SD = 1.00, range = 0-2). The decision to exclude participants in one condition did not exclude all low-performing children, but rather attempted to eliminate the possible confound of some children having received an inadequate treatment schedule.

Participants

Thirteen kindergarten children with SLI (*M* = 5;5, *SD* = 0;4, range 5;0 – 6;1) were recruited through language screenings. Fifty-four percent of participants were boys and 46% were girls. Participants had the following characteristics: 85% White-Non-Hispanic, 8% Black /African American-Non-Hispanic and 8% White with ethnicity not reported. In terms of parent characteristics, 54% of parents were married, 23% were single, and 23% were divorced. In terms of mother's education, 38% had partial college, 23% were high school graduates, 15% were college graduates, 15% had partial high school, and 8% had a graduate degree. In terms of father's education, 38% had partial college, 23% were college graduates, 23% were not reported, 8% had a graduate degree, and 8% were high school graduates.

Table 1 contains a summary of children's results on all pretreatment test measures. Percentile scores are provided for easier comparison of subtest scaled scores and composite scores, which use different scales. Eligible children were required to (1) be enrolled in or eligible for kindergarten; (2) pass a hearing screening (ASHA, 1997); (3) score at or above the 16th percentile for nonverbal cognition on the Reynolds Intellectual Assessment Scale (Reynolds & Kamphaus, 2003); (4) have a Core Language Score at or below the 10th percentile on the CELF–4th edition (Semel et al., 2003); and (5) score at or below the 10th percentile on at least one of two vocabulary measures: the DELV Semantic Subtest (Seymour et al., 2005) or the CELF-4 Word Classes Subtest. For vocabulary scores, most children (46%) qualified on both the DELV and the CELF, 38% qualified on the DELV only, and 15% qualified on the CELF only. Children also completed supplementary tests to further characterize their abilities, including: the CTOPP 2nd edition Elision, Sound Matching, Blending Words, Nonword Repetition, and Memory for Digits Subtests (Wagner et al., 2013); the CELF Understanding Spoken Paragraphs Subtest; and the Goldman-Fristoe Test of Articulation 2nd edition (Goldman & Fristoe, 2000).

Table 1

Percentile scores for participants on standardized clinical tests

Test	Mean	SD	Range	% at or below 10th percentile
RIAS[1] Nonverbal IQ	57	16	34-95	0%
CELF[2] Core Language	4	3	<0.1-10	100%
Vocabulary: DELV[3] Semantic	9	11	1-37	85%
Vocabulary: CELF[2] Word Classes	19	24	2-75	62%
CELF[2, 4] Concepts & Following Directions	9	11	0.1-37	77%
CELF[2, 4] Word Structure	8	7	1-25	77%
CELF[2, 4] Recalling Sentences	11	15	0.1-50	77%
CELF[2, 4] Formulating Sentences	8	7	0.1-25	77%
CELF[2] Understanding Spoken Paragraphs	11	9	1-25	69%
CTOPP[5, 6] Nonword Repetition	18	22	2-75	69%
CTOPP[5] Phonological Memory	14	19	<1-68	62%
CTOPP[5] Phonological Awareness	6	5	<1-14	77%
GFTA[7]	27	20	2-58	15%

[1]Reynolds Intellectual Assessment Scale (Reynolds & Kamphaus, 2003); [2]Clinical Evaluation of Language Fundamentals 4th edition (Semel et al., 2003); [3]Diagnostic Evaluation of Language Variation (Seymour et al., 2005); [4]Scores on this subtest contributed to the CELF Core Language score; [5]Comprehensive Test of Phonological Processing 2nd edition (Wagner et al., 2013); [6]Scores on this subtest contributed to the CTOPP Phonological Memory composite score; [7]Goldman-Fristoe Test of Articulation 2nd edition (Goldman & Fristoe, 2000).

Treatment

The book reading treatment was based on the procedures used by Justice and colleagues (2005) in a clinical trial with kindergarteners from low-income homes. In that study, the majority of children (77%) made meaningful gains, defined by an improvement of at least 4 points from pre- to post-treatment vocabulary testing of treated words. For participants with low vocabulary, Justice and colleagues found a large effect size ($d = 1.34$) for the elaborated book reading treatment condition compared to a control condition with no treatment.

Treatment materials. Justice and colleagues (2005) selected 60 words from 10 commercially available storybooks following the guidelines of Beck, McKeown, and Kucan (2002). All words were "tier two" words, which are frequently used by mature language users and which appear across many academic contexts. Children in this study were taught one set of 30 words (labeled set A or set B) in each of the two treatment conditions to which they were assigned, for 60 words total across the entire treatment. Thus, of the 13 participants whose results are reported here, 4 learned set A words and 9 learned set B words in their first treatment condition.

The storybooks, words, and word characteristics for set A and set B are listed in Appendix B. The 60 treatment words included nouns (n = 16), verbs (n = 25), and adjectives (n = 19). For the current study, the 10 books were divided into two sets of 5 books each (i.e., 30 words each) that were matched on word type: set A contained 9 nouns, 12 verbs, and 9 adjectives, while set B contained 7 nouns, 13 verbs, and 10 adjectives. A chi-square test for independence was performed to examine the composition of nouns, verbs, and adjectives across the two word sets. The number of nouns, verbs, and adjectives was not significantly different between word sets, $X^2(1) = 2.00$, $p = 0.20$.

Four variables were computed from the online child calculator by Storkel and Hoover (2010): word length in phones, positional segment frequency average, biphone frequency average, and number of neighbors. Word length in phones describes the number of sounds in each word. Word length was not a variable of interest, but was analyzed to check the parity between word sets A and B. The positional segment frequency average is a measure of phonotactic probability found by calculating the *positional segment sum*—the sum of the log frequencies of all words in the database containing the target phoneme in the same position as in the target word and dividing by the sum of the log frequencies of all words in the database that had any phoneme in the target word position (Storkel, 2004b). The positional segment sum is then divided by the word length in phones to compute the positional segment frequency average. The biphone frequency average is another measure of phonotactic probability found by calculating the *biphone sum*—the sum of the log frequencies of all words in the database containing the target phoneme pair in the target word position and dividing by the sum of the log frequencies of all words in the database that had any phoneme in the target word position (Storkel, 2004b). The biphone sum is then divided by the word length in phonemes to compute the biphone frequency average. For positional segment frequency and biphone frequency averages, a higher value indicates a more likely combination of sounds. The number of neighbors is a measure of neighborhood density calculated by counting the number of words in the database that differed by one phoneme (either added, deleted, or substituted) in any position from the target word. The more neighbors a word has, the denser its phonological neighborhood is said to be. Individual values and summary characteristics across word sets A and B are included in Appendix A. Independent samples t-tests were performed to compare values for word length in phones, positional segment frequency average, biphone frequency average, and number of neighbors across word sets. The comparison for word length in phones approached significance, with set A words (M = 4.83, SD = 1.26) slightly longer than set B words (M = 4.30, SD = 1.09), $t(58) = 1.75$, $p = 0.09$. All other comparisons were not significant, $t\ (58) < 0.32$, $p > 0.58$.

Treatment form. Treatment sessions were conducted by research assistants one-on-one with the child in a quiet room in his or her school or another convenient location (e.g., the local library). Readings and related activities were completed for two books per session, with each session lasting 20-30 minutes. For each book, treatment consisted of: (1) a preview, in which the treatment provider showed the child a picture associated with each of the 6 words in the upcoming storybook and read a definition and synonym for each word; (2) the reading of the storybook, with the treatment provider pausing to read the synonym after each word occurred in the text; and (3) a review, in which the treatment provider showed the child another related picture for each word and read a context sentence and the definition. This is similar to the elaborative procedures used by Justice and colleagues (2005), with the addition of the synonym. Treatment providers conducted the sessions following printed scripts.

Two treatment sessions occurred per week for approximately 7.5 to 11.5 weeks for the balanced and maximize dose frequency regimens, respectively (depending on child attendance). For children in the balanced condition, each of the 6 books were read on 6 different occasions, and children heard 6 exposures to the words during each reading. For children in the maximize dose frequency regimen, each of the 6 books were read on 9 different occasions, and children heard 4 exposures to the words during each reading (achieved by eliminating the synonym during the storybook reading and the definition during the review). Thus, at the end of treatment, all children had accumulated 36 exposures to each of the 30 treated words.

Treatment fidelity. For 20% of sessions, an observer watched the videotaped session and completed a checklist to ensure that each target word and its associated definition, synonym, and context were read accurately and the correct number of times. From this checklist, the observer tallied the total number of exposures to each word administered and divided this score by the prescribed number of exposures. This score was 99.88%, indicating that the number of exposures was administered with high fidelity. Next, the observer tallied the total number of treatment forms (i.e., definitions, context sentences, and synonyms) correctly administered and divided this number by the number of possible correct treatment forms. This score was 99.83%, indicating that the treatment form was administered with high fidelity.

Outcome Measure

The primary outcome measure was a definition task, which was administered before treatment, immediately following the conclusion of treatment, and 2 weeks following the conclusion of treatment. On each of these three occasions, the task was administered across two sessions with 15 set A and 15 set B words being tested in each session. The children first heard three practice words (bed, ball, candy), which were words likely to be known by children with SLI. Then, set A (n = 15) and set B (n = 15) words were presented in random order along with familiar words (n = 10, e.g., chair, teacher, apple). For each word, the child received the prompt, "Tell me what [word] means." Prompts were pre-recorded and presented with computer software so that pronunciations of the words were consistent across tasks. Children's responses were audio recorded and transcribed for later scoring.

The definition scoring procedures from Study 1 (Storkel et al., 2017) were also used in Study 2 (the current study). In Study 1, the principal investigator, project coordinator, and 5 graduate research assistants consulted dictionaries to create a scoring rubric for each word that listed elements of an accurate and complete definition (e.g., decided = make up + mind; chose) following the recommendations of McGregor, Oleson, Bahnsen, and Duff (2013). Possible scores included: 0 points for an incorrect or absent definition (e.g., decided = way to go), 1 point for an appropriate use of the word in a sentence (e.g., decided = decided what way to go) or for a vague definition (e.g., decided = think), 2 points for a conventional definition containing at least one critical element but lacking other critical elements (e.g., decided = think what way to go, missing element of choosing), and 3 points for a complete and accurate definition including all critical elements (e.g., decided = make up your mind). Each response was independently scored by two judges following the rubric guidelines. These two scores were compared and disagreements were resolved by consensus. In rare cases when the two raters could not reach consensus, they consulted a third judge.

For the analyses, children's definitions scored as 2 or 3 (i.e., at least a partially accurate definition) were considered correct and definitions scored as 0 or 1 were considered incorrect. The analyses for this study focus only on the post-test conducted immediately following treatment because this test represents immediate learning from the treatment while the 2-week post-test represents extended retention after treatment was withdrawn.

Independent Variables

Fast mapping ability. Fast mapping describes the task of quickly matching an unknown word to its referent or meaning. Our measure of fast mapping ability was the raw score for the Fast Mapping Novel Verbs subsection of the DELV Semantics Subtest. These test items require the child to answer questions about derived word forms by pointing to an image. The child is told that the examiner made up some silly words. Verb items are nonwords with a unique meaning. For example, the child is shown a picture of a novel action as the examiner reads: "The man is lelling the clown…Which one is the leller?" (Seymour et al., 2005, Record Form p. 20). The test contains 12 items assessing fast mapping of novel verbs. Norms are not available for these subtest items alone, so raw scores were used for all children.

Phonological working memory. Phonological working memory describes a child's ability to store and manipulate sound sequences. Our measure of phonological working memory was the Phonological Memory (PM) composite score of the CTOPP-2 and the Nonword Repetition subtest score. The PM composite score is derived from two subtests: Memory for Digits, which requires children to repeat strings of digits presented via CD; and Nonword Repetition, which requires children to repeat nonwords of increasing length presented via CD. Nonword repetition and digit span tasks are frequently used as measures of phonological working memory in studies of children with SLI. We used both the composite score of these two standard scores (M = 100, SD = 15) and the Nonword Repetition subtest standard score (M = 10, SD = 3) for the analyses

Semantics. Our two measures of semantics were the Word Classes Subtest of the CELF-4 and the DELV Semantics Subtest.

CELF-4 Word Classes Subtest. The Word Classes Subtest requires children to identify two words presented in a field of 3 or 4 that go together (e.g., Helicopter, birdhouse, kite. Which two go together?) and then to state why the two items go together. For example, one subtest item asks "Helicopter, birdhouse, kite…How do the words _____ and _____ go together?" (Semel et al., 2003, Record Form 1 p. 9). This assesses the depth of the child's knowledge of the meaning of the word and his or her ability to identify semantic relationships between words (i.e., categories, functions, attributes, or other similarities they share). We used the subtest standard score for the analyses (M = 10, SD = 3).

DELV Semantics Subtest. The DELV Semantics Subtest contains items assessing verb contrasts (e.g., "The man isn't *walking*, he's…crawling"; Seymour et al., 2005, Record Form p. 15); preposition contrasts (e.g., "She's not looking at the radio, she's listening…to the radio"; Seymour et al., 2005, Record Form p. 17); quantifiers (e.g., "Is every man riding a horse?"; Seymour et al., 2005, Record Form p. 18); fast mapping with real verbs (e.g., "The postal worker is handing the letter to the boy…Which one was the hander?"; Seymour et al., 2005, Record Form p. 29); and fast mapping with novel verbs (described previously). We used the subtest adjusted standard score for the analyses (M = 10, SD = 3). This score was adjusted for the child's parents' level of education, following the DELV guidelines.

Language ability. Our two measures of language ability were the CELF-4 Core Language Score and the CELF-4 Understanding Spoken Paragraphs Subtest.

CELF Core Language. This composite score is derived from four subtest scores: Concepts & Following directions, which requires children to point to images following directions involving linguistic concepts (e.g., "After you point to the shoe, point to the fish"; Semel et al., 2003, Stimulus Book 1 Item 5); Formulated Sentences, which requires children to make up a grammatically correct sentence about a picture using a given word; Recalling Sentences, which requires children to repeat sentences of increasing length and complexity; and Word Structure, which tests children's knowledge of grammatical morphemes. The combination of these four tests assesses children's overall receptive and expressive language ability. We used the composite Core Language Score derived from the combination of these four subtest scores for the analyses (M = 100, SD = 15).

CELF Understanding Spoken Paragraphs. The supplemental Understanding Spoken Paragraphs subtest assesses children's ability to answer content and critical thinking questions after listening to increasingly longer and more complex paragraphs. After listening to each of three short narratives, children are asked questions that require them to recall details from the story (e.g., "What did the brothers hear before they went to bed?"; Semel et al., 2003, Record Form 1 p. 14) or to use critical thinking to arrive at answers that were not explicitly stated (e.g., "What do you think Rudy and Louis will do now that the rain has stopped?"; Semel et al., 2003, Record Form 1 p. 14). This supplemental test of receptive language was included because it was thought to resemble the task that children faced during the treatment. That is, children needed to pick out words and their meanings from longer discourse and recall those details. We used the subtest adjusted standard score for the analyses ($M = 10$, $SD = 3$).

Phonotactic probability. As described previously, phonotactic probability measures how likely a particular sound sequence is to occur. It is represented by two numbers, the positional segment frequency average and the biphone frequency average (details on the calculation of these two numbers were described in the Treatment Materials section above). For positional segment frequency and biphone frequency averages, a higher value indicates a more likely combination of sounds. Raw values for both numbers were used in the analyses.

Neighborhood density. As described previously, neighborhood density describes the number of words that are similar to a word by the addition, deletion, or substitution of one sound. The more phonological neighbors a word has, the denser its phonological neighborhood is said to be. The raw number of phonological neighbors were used in the analyses.

Part of speech. Part of speech (noun, verb, adjective) for each treatment word was listed and verified by graduate research assistants. In cases where a word could function as a different part of speech depending on the context, the part of speech was determined from the use in the storybook (e.g., *gulp* is listed as a noun due to its context, "in one gulp").

Results

Child Characteristics

The first research question was: Do child characteristics (specifically, fast mapping ability, phonological working memory, semantics, and language ability) relate to the number of words learned during treatment? To investigate this, simple correlations were conducted comparing each independent variable (i.e., DELV fast-mapping scores, CTOPP Nonword Repetition scores, CELF and DELV vocabulary scores, and CELF Core Language and Understanding Spoken Paragraphs scores) to the dependent variable (i.e., the number of words correctly defined by children at the immediate post-test). Results are reported below in Table 2 with the data in the first numbered column addressing the research question. The columns numbered 2 through 7 show the correlations among measures. Significant correlations are marked. DELV fast-mapping scores, $r(13) = .63$, $p < .05$ and CELF Understanding Spoken Paragraph (USP) scores $r(13) = .61$, $p < .05$ were significantly positively correlated with treatment outcomes. Both of these correlations show a large effect size (i.e., $r \geq .10$ for small, $r \geq .30$ for moderate, $r \geq .50$ for large; Cohen, 1988).

Table 2

Correlations among treatment outcomes and child characteristics variables

	1	2	3	4	5	6	7	8
1. Number of Words Correctly Defined After Treatment	1.00							
2. DELV Fast Mapping Novel Verbs Raw Score	0.63*	1.00						
3. CTOPP Nonword Repetition Scaled Score	-0.17	-0.32	1.00					
4. CTOPP Phonological Memory Composite Score	-0.19	-0.39	0.90**	1.00				
5. CELF Word Classes Scaled Score	0.17	-0.30	0.18	0.13	1.00			
6. DELV Semantics Scaled Score	0.49	0.48	-0.55	-0.51	-0.18	1.00		
7. CELF Core Language Standard Score	0.52	0.24	0.14	0.36	0.22	0.38	1.00	
8. CELF Understanding Spoken Paragraphs Scaled Score	0.61*	0.28	0.25	0.21	0.33	0.50	0.71**	1.00

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

To further examine the significant relationships, scatterplots were created for each significant correlation. Figure 1 plots children's raw scores on the DELV fast mapping novel verbs items against the number of words correctly defined after treatment. As scores on the DELV fast mapping items increased, the number of words learned increased as well, suggesting that children who performed better on the fast mapping task learned more words from interactive book reading. The r^2 value indicates that 39% of the variance in the number of words learned was accounted for by the DELV fast mapping score. However, it is important to note that the data may be influenced by several high-scoring outliers, which in such a small sample may have unduly influenced the correlation.

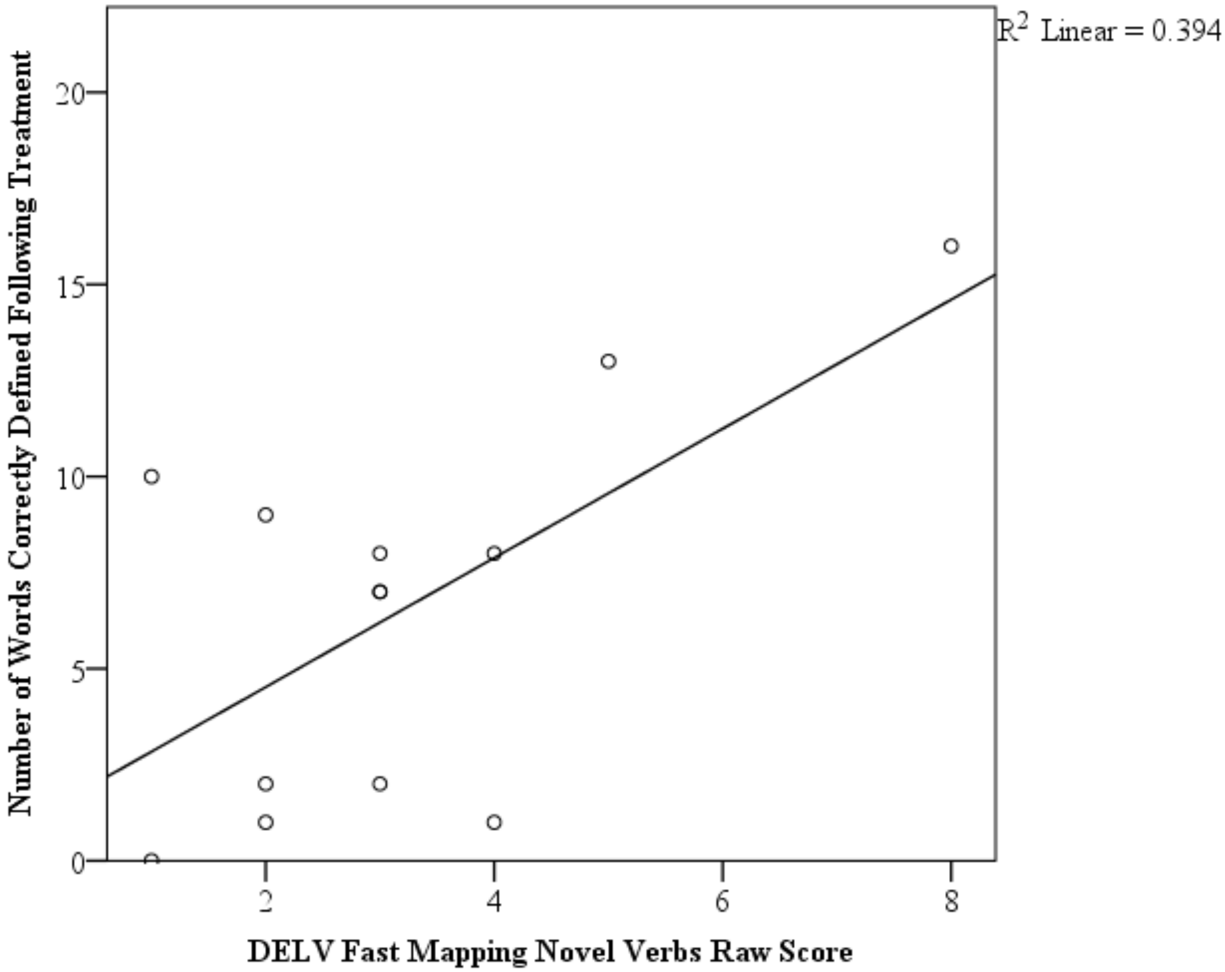

Figure 1. Scatterplot illustrating the relationship between the number of words learned in treatment and children's DELV Fast Mapping Novel Verbs raw score.

Figure 2 plots scaled scores on the CELF USP subtest against the number of words correctly defined at post-treatment. As scores increased on USP, the number of words learned also increased. The r^2 value indicates that 37% of the variance in the number of words learned was accounted for by the CELF USP score. This suggests that children who were better able to understand short narratives also were better able to learn words during interactive book reading. However, it is also possible that other skills might account for this relationship. For example, CELF USP was significantly correlated with CELF Core Language Score. This suggests that the CELF USP might be tapping general language ability and that general language ability is related to word learning during interactive book reading rather than the specific ability to understand a short narrative.

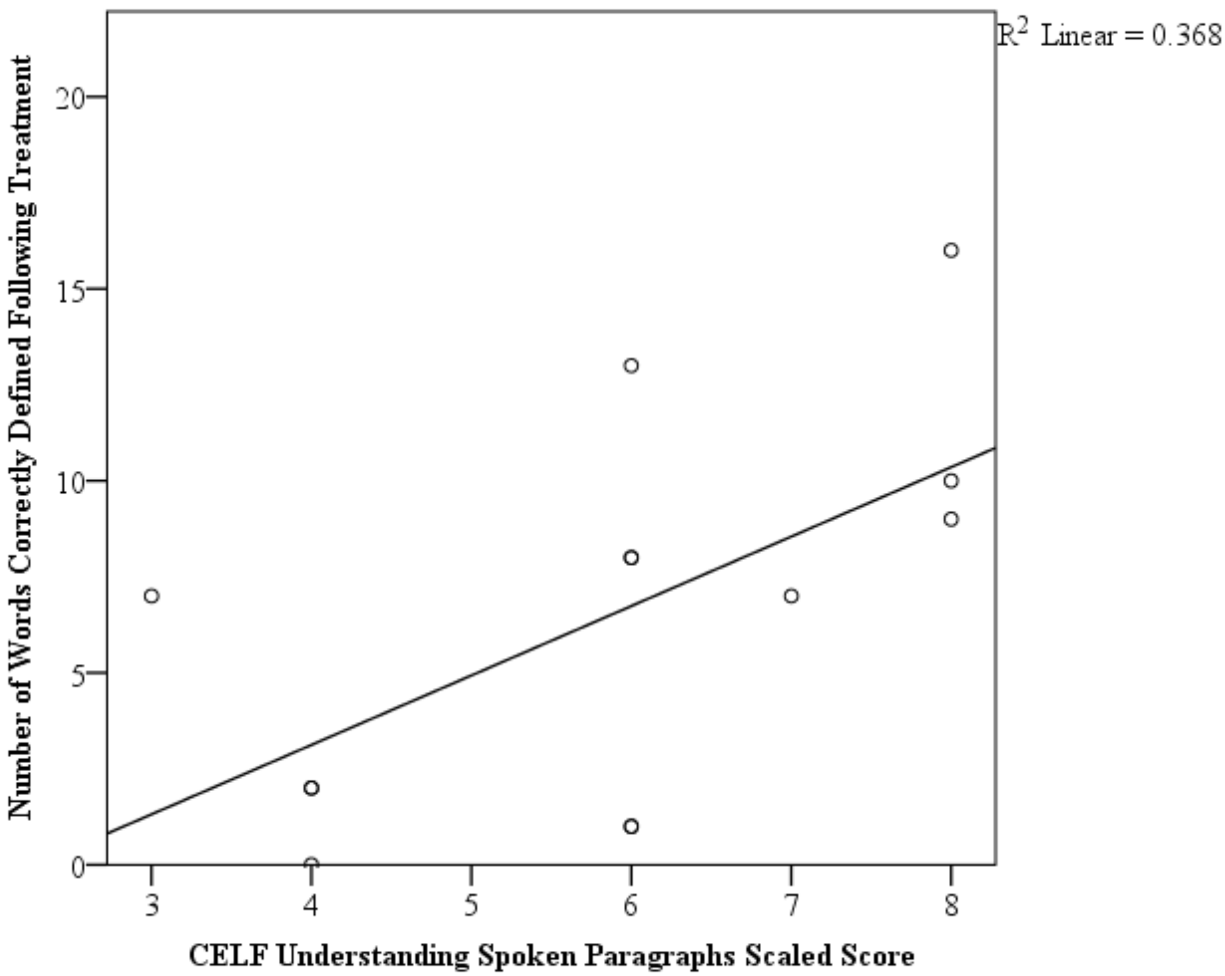

Figure 2. Scatterplot illustrating the relationship between the number of words learned in treatment and children's CELF Understanding Spoken Paragraphs scaled score.

A multiple linear regression was calculated to examine the relationship between the two significant child characteristics variables (DELV fast mapping raw scores and CELF USP scaled scores) with the number of words that children correctly defined following treatment. A process of model building was completed to examine all possible effects. The results are reported below in Table 3. In the first model, the DELV fast mapping was added first, accounting for 39% of the variance in words learned. Next, the CELF USP was added to create model 2, which accounted for 60% of the variance in words learned (i.e., the R square value increased by 0.20). The F change going from model 1 to model 2 was significant, F change (1, 10) = 5.07, $p < 0.05$. Next, the variables were added in the opposite order. In model 3, the CELF USP was added first, accounting for 37% of the variance in words learned. Next, the DELV fast mapping was added to create model 4, which was identical to model 2.

Table 3

Multiple linear regression results for significant child characteristics variables

Variable	Model 1	Model 2	Model 3	Model 4
DELV Fast Mapping Novel Verbs Raw Scores	1.68* (0.63)	1.33* (0.56)		1.33* (0.56)
CELF Understanding Spoken Paragraphs Scaled Scores		1.40* (0.62)	1.81* (0.71)	1.40* (0.62)
F	7.14	7.43	6.41	7.43
p	0.02	0.01	0.03	0.01
R^2	0.39	0.60	0.37	0.60

* $p < 0.05$, ** $p < 0.01$

Partial correlations were also computed for each variable. Results showed that the DELV fast mapping score accounted for 36% of the variance in words learned when controlling for CELF USP scores, $p < .05$. Conversely, the CELF USP scores accounted for 34% of the variance in words learned when controlling for DELV fast mapping scores, $p < .05$. Overall, these results suggest that DELV fast mapping scores and CELF USP scaled scores uniquely predicted variance in the number of words learned. That is, both variables seem to tap different underlying abilities that influence the number of words learned.

Word Characteristics

The second research question was: Do word characteristics (specifically, phonotactic probability, neighborhood density, and part of speech) relate to the percentage of children that learned each word? To investigate the relationship between phonotactic probability, neighborhood density, and learning for each word, simple correlations were conducted comparing each measure (i.e., positional average, biphone average, number of neighbors) to the percentage of children who learned each word. Results are reported below in Table 3, with correlations of interest listed in the first numbered column. Columns numbered 2 through 4 show the correlations among measures. Since only one correlation of interest was significant, a regression analysis was not conducted. The number of phonological neighbors $r(60) = .61$, $p < .05$ was significantly negatively correlated with treatment outcomes. This correlation shows a small effect size.

Table 4

Correlations among treatment outcomes, phonotactic probability and neighborhood density

	1	2	3	4
1. % of Children Who Learned Treatment Word	1.00			
2. Positional Average	-0.09	1.00		
3. Biphone Average	-0.11	.607**	1.00	
4. Number of Neighbors	-0.29*	.362**	.331**	1.00

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

A scatterplot illustrating the data for the correlation between words learned and number of neighbors is included in Figure 3. Words with from sparser phonological neighborhoods were generally correctly defined by more children. The r^2 value indicates that only 8% of the variance in the percentage of children who learned each word was accounted for by neighborhood density. The data are also somewhat suppressed since many words were not learned by any children. The range of the data for number of neighbors also does not appear to evenly cover the possible spectrum; most words had 10 or fewer neighbors, with relatively few words from more dense neighborhoods. This likely occurred because the stimuli were real words which were typically longer than the specifically selected nonwords used in many studies of initial word learning.

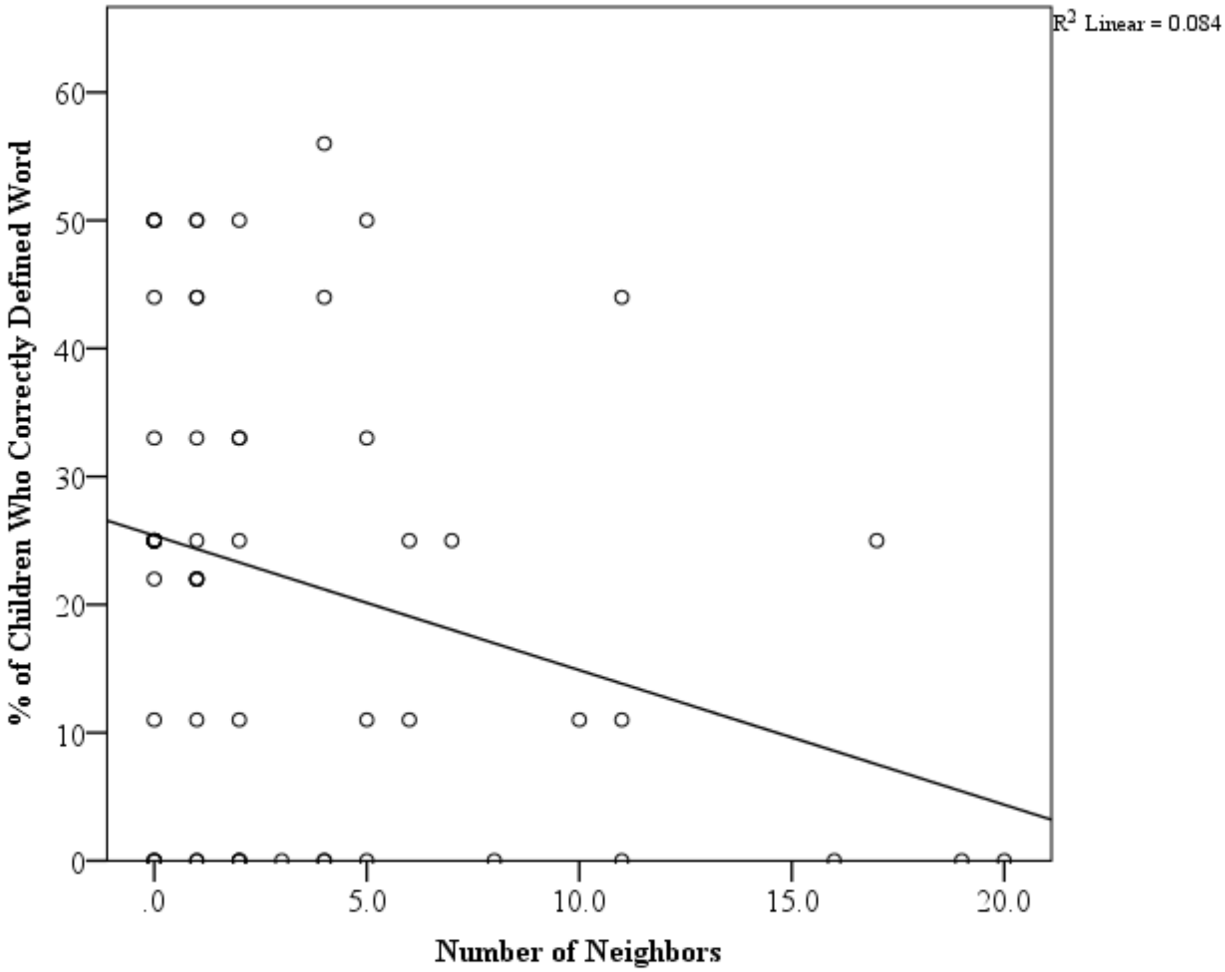

Figure 3. Scatterplot illustrating the relationship between the percentage of children who correctly defined each word following treatment and the number of phonological neighbors for each word.

To investigate the relationship between part of speech and learning for each word, a one-way independent samples ANOVA was conducted. The ANOVA compared part of speech (3 levels: noun, verb, and adjective) to the percentage of children who learned each word. There was no statistically significant effect of part of speech, $F(2, 57) = 0.41$, $p = 0.67$, $\eta^2 = .01$. The effect size was small (i.e., $\eta^2 \geq .01$ for small; Cohen, 1988). Table 4 contains a summary of the mean percentage of children who learned words of each type. Descriptively, it appears that children showed the best learning for nouns, followed by adjectives, with the lowest learning for verbs. This aligns with the results found in Study 1 (Storkel et al., 2017).

Table 5

Percentage of children who learned each word by part of speech

	N	Mean	SD
Noun	16	24.38	19.66
Adjective	19	22.58	19.86
Verb	25	19.36	15.40
Total	60	21.72	17.88

Discussion

Two research questions were addressed in the current study: (a) do child characteristics (specifically, fast-mapping ability, phonological working memory, semantics, and language ability) relate to the number of words learned during treatment? and (b) do word characteristics (specifically, phonotactic probability, neighborhood density, and part of speech) relate to the percentage of children that learned each word? Regarding the first research question, results indicated that children who learned more words in treatment generally had higher fast mapping ability (i.e., DELV fast mapping novel verbs raw scores) and higher receptive language ability (i.e., CELF Understanding Spoken Paragraphs scaled scores). Regarding the second research question, results indicated that children showed better learning for words with fewer phonological neighbors (i.e., words from sparse neighborhoods).

DELV Semantics scores significantly predicted word learning outcomes in Study 1 (Storkel et al., 2017). Standardized measures of vocabulary have not been consistently predictive of children's word learning in studies on the initial word learning of children with SLI (Gray, 2003, 2006; Kiernan & Gray, 1998; Rice et al.,1990; Rice, Oetting, Marquis, Bode, & Pae, 1994) and in the study of typically developing kindergarteners from low-income homes on which the current interactive book reading treatment is based (Justice et al., 2005). Storkel and colleagues posited that the DELV's items focusing on fast mapping may have explained this finding. Although the overall DELV Semantics score was not significantly correlated to treatment outcomes in Study 2, the raw number of fast mapping novel verbs items was significantly positively correlated to the number of words children learned. The DELV's authors argued that these items should be less influenced by children's previous knowledge than measures of vocabulary size because they test the ability to learn new words in context (Seymour et al., 2005). The DELV fast mapping items are also similar to the task children faced in the vocabulary treatment. Elaborated exposures to the words included context sentences from the storybooks and additional context sentences provided after each reading. Children who made better use of this input by mapping meanings onto new words from context would have gained additional information about the words' meaning from each book reading session.

Although the significant findings relating DELV performance to treatment outcomes in both Studies 1 and 2 are a promising indication for the predictive ability of this measure, there are some limitations to the current finding. Study 2 involved a small sample of children, and these effects may change as more participants are added, especially considering that several outlying cases may be driving this relationship. However, given the nature of the DELV tasks and these preliminary findings, further study of the DELV or similar fast mapping measures and their predictive value for word learning treatment outcomes is warranted.

Receptive language ability, measured by the CELF Understanding Spoken Paragraphs (USP) subtest, also significantly predicted children's treatment outcomes. This supports Kan and

Windsor's (2010) finding that children with SLI and lower receptive language had poorer word learning abilities than their peers. Like the DELV fast mapping items, the CELF USP items also resemble the task faced by children in the interactive book reading treatment. In both, children must listen to longer units of discourse and later recall details from that discourse. The CELF Core Language Score was significantly correlated with the CELF USP score (which is a supplemental test not included in the Core Language Score). This suggests that overall language ability, not the specific skill of answering questions about short narratives, may predict treatment outcomes. However, overall CELF Core Language Scores did not significantly predict treatment outcomes. Several of the subtests included in the Core Language Score are more expressive in nature (e.g., repeating sentences, formulating sentences about a picture). Therefore, it is possible that it is children's receptive language ability specifically, rather than their general language ability, that predicts treatment outcomes. A regression analysis showed that both fast mapping and receptive language explained unique variance in the number of words learned, suggesting that these two tests measured different skills and that both are important in learning words from an interactive book reading treatment.

Regarding word characteristics, results suggested that words with more phonological neighbors were typically learned by fewer children. Experimental literature (e.g., Storkel, Armbruster, & Hogan, 2006) has indicated that typically developing children show better learning for words from sparse neighborhoods (i.e., words with fewer phonological neighbors), likely because these words are distinct from the words already in the lexicon, aiding in the creation of a new representation. However, research by Storkel and Lee (2011) found improved later retention for words from dense neighborhoods, suggesting that integrating a new word with many similar words may aid children with creating a representation in the process of extended mapping. Preliminary research on children with SLI by Gray and colleagues (2014) indicated that children may benefit from sparse neighborhoods, possibly due to their creating more holistic representations. Studies 1 and 2 are the first to examine the long-term word retention of children with SLI learning real words through interactive book reading, and this finding warrants replication from further studies. The current stimuli were not chosen for neighborhood density, so the range of values for number of neighbors does not evenly cover the spectrum. That is, the word set was mostly clustered at 10 and fewer neighbors with few words from dense neighborhoods, possibly due to word length.

However, it is important to note the difficulty of choosing a set of real words that are tightly controlled for one variable. The words in the current study were chosen primarily because they were used in the selected storybooks and because they were words likely to be useful for children across many everyday contexts. Experimental studies, which tend to teach nonwords of controlled length, do not have to account for these considerations; this makes controlling for specific features of words and finding significant effects of these features on learning more straightforward. Several possible confounding features of the words used in this study may have impacted the results. As previously noted, the spread of neighborhood density values was not equal. This is not easily remedied, however, as other factors tend to vary with neighborhood density. Words from dense neighborhoods tend to be shorter and more common than words from sparse neighborhoods (Storkel, 2004a). A word's phonotactic probability or neighborhood density may also be influenced by the presence of common affixes and suffixes (Storkel, 2004b). Other factors not quantified in this study may be important as well. At least when considering the age when children acquire words, there seems to be an advantage for words that are more imageable (i.e., easily pictured) or more concrete (i.e., more physical than abstract; Bird, Franklin, & Howard, 2001). Factors like these are more difficult to quantify. Given these challenges, it is perhaps not surprising that neighborhood density explained such a small amount of the variance in words learned in this study. The difficulty of controlling sets of real words for word characteristics does not make such investigations less valuable. However, in designing such studies one must consider how to code possible variables of interest, as well as how to achieve enough statistical power to find small differences in noisy data. Increasing the effectiveness of the treatment so that children learned more words would also strengthen such analyses.

Study 1 vs. 2 Discrepancies

Some of the factors hypothesized to correlate with children's treatment outcomes were not significant in Study 2 or in Study 1. The measure of overall language, the CELF Core Language Score, was not significant in either study. It is possible that a broad omnibus measure of language does not capture a specific enough set of skills to explain treatment outcomes. The CELF Word Classes subtest, one of the two measures of semantics, was also not significant in Study 2 or Study 1. It was thought that since this measure tested children's understanding of word relationships that it may have been a deeper measure of semantics, superior to tests that ask children to only identify words. However, the outcome measure for the study asked children only to define the words, not to use any metalinguistic skills to compare them to one another. This could be why the DELV Semantics subtest, which focuses more on mapping meanings to words, was a stronger predictor of outcomes in both studies. For word characteristics, both studies found no significant effect for phonotactic probability and a descriptive but not significant difference for part of speech. In both studies, overall learning for the words showed a floor effect, with many words not learned by any children (i.e., 30% of words in Study 1 and 28% of words in Study 2). This floor effect, combined with the fact that words were not specifically chosen for their characteristics, may make it difficult to detect small effects of word characteristics that may be present.

The DELV was a significant predictor of children's word learning in both Study 2 and Study 1, but different scores were significant in each study. In Study 1, the DELV Semantics subtest significantly predicted the number of words learned. In Study 2, the DELV Semantics subtest was not a significant predictor of words learned, but the raw number of fast mapping novel verbs items from the Semantics subtest was. Since the raw fast mapping scores were not included in Study 1, this result cannot be compared across studies. However, it is possible that these items are a better predictor of word learning outcomes than the Semantics subtest in general.

Other variables predicted outcomes in one study but not the other. The CTOPP Nonword Repetition score significantly predicted children's outcomes in Study 1 but the CTOPP Nonword Repetition and Phonological Memory Composite scores did not significantly predict outcomes in Study 2. The measure of receptive language, the CELF USP subtest score, was significantly correlated with the number of words learned in Study 2 but not in Study 1. Children received a variety of different exposures (i.e., 24, 36, or 48) in Study 1, potentially tapping into weaknesses and leading semantic and phonological processing measures (i.e., DELV Semantics, CTOPP Nonword Repetition) to be the significant predictors of word learning. In Study 2, kids received the adequate number of exposures and that is minimizing the contribution of phonological processing (CTOPP Nonword Repetition) and allowing other higher-level weaknesses to be identified (CELF USP). Finally, neighborhood density was significant in Study 2 but not in Study 1. As described previously, many words were rarely or never learned in each study, words were not chosen for their neighborhood density values, and it is difficult to carefully control for one variable when teaching real words. This may result in small effects that are difficult to detect and replicate across samples of children.

Implications and Future Directions

The factors that significantly predicted children's word learning in this study suggest that further study is needed to determine whether some children would benefit from modifications the treatment paradigm. Children's fast mapping and receptive language abilities significantly predicted how many words they learned. Children who struggle with mapping meanings to words may benefit from more explicit, less contextual teaching of the words. In the naming task during treatment, some children would reliably repeat part of the word's associated context sentence, but not the word itself. It was not clear if these children understood that one word embedded within the context sentence was the word to be learned. It may help to draw these children's attention to the target word. For example, during the first few sessions, the experimenter could say "Our word is marvel. I'm going to tell you what marvel means. First, you say marvel." This may help children identify the word to be learned and make more sense of the input surrounding the word.

In addition, treatment providers could add contingent feedback to sessions to correct children's understanding of word meanings and to aid in accurate mapping of meanings to new words. In the current treatment paradigm, children responded to a prompt and then heard the definition, but did not receive targeted feedback. For example, the definition picture presented during treatment for the word "swift" was a cheetah. When asked to "Tell me what swift means," children would sometimes respond "a cheetah." At that point, the treatment provider could explain that a cheetah is an example of something that is swift, but that the word actually means fast, and many things other than cheetahs might be swift as well. In this way, children's early misperceptions about the word's meanings could be addressed in a targeted way.

If children struggle with the receptive language task of answering detail and inferencing questions from stories, it is not clear how much information they receive from the input during interactive storybook reading sessions. This could be examined in future studies by asking children story comprehension questions after each reading. It is possible that the storybooks used were difficult for some children to understand, making the exposures received during book reading less useful. If children struggle to answer questions about the stories, they may benefit from simplified stories with simpler syntax, more frequent words, or a streamlined narrative. Children with lower receptive language may also benefit from receiving the treatment in stages, with explicit teaching of the definitions in the first phase and more contextual teaching in storybooks only after children showed some learning for the target words.

This study suggested that children may show poorer learning for words from denser phonological neighborhoods (e.g., 10 or more neighbors). For these words, children may benefit from explicit teaching of the phonological form. The current treatment paradigm did not include explicit training on the word's phonology. To draw children's attention to the word's form, the treatment provider could say, "Our word is marvel. Let's clap the syllables together. Marvel starts with the /m/ sound. What are some other words that start with /m/?" Adding phonological teaching is supported by findings by Gray (2004), who showed that children with SLI have more difficulty with phonology than semantics in initial word learning. Future studies could test different phonological cues and their influence on children's learning of words from sparse or dense phonological neighborhoods.

Conclusion

This study provided an initial investigation of the child and word characteristics that influenced the word learning outcomes of children with SLI in an interactive book reading treatment. Results indicated that children with better fast mapping and receptive language abilities may learn more words from an interactive book reading treatment. Children with SLI also may show better learning for words with fewer phonological neighbors. These findings provide areas of interest for additional, larger studies. Investigating the factors that improve word learning outcomes for children with SLI could help clinicians improve vocabulary treatment outcomes for these children. Clinicians could choose treatments and tailor treatment methods to specific children's strengths and weaknesses based on their profiles on clinical tests. Specific teaching methods could be used for words with certain phonological characteristics (e.g., words from dense neighborhoods). These promising predictors support and expand on the experimental literature on word learning in children with SLI. This study's findings contribute to the initial work in translating experimental literature into practice recommendations.

References

Alt, M. (2011). Phonological working memory impairments in children with specific language impairment: Where does the problem lie? *Journal of Communication Disorders, 44*(2), 173-185. doi:http://dx.doi.org.www2.lib.ku.edu/10.1016/j.jcomdis.2010.09.003

Alt, M., & Plante, E. (2006). Factors that influence lexical and semantic fast mapping of young children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 49*(5), 941-954. doi:http://dx.doi.org.www2.lib.ku.edu/10.1044/1092-4388(2006/068)

ASHA. (1997). Guidelines for screening for hearing impairment-school-age children, 5 through 18 years. *ASHA, 4,* IV-74ff-IV-74hh.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life.* New York: Guilford Press.

Beverly, B. L., & Estis, J. M. (2003). Fast mapping deficits during disambiguation in children with specific language impairment. *Journal of Speech-Language Pathology and Audiology, 27*(3), 163-171.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments & Computers, 33*(1), 73-79.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Estis, J. M., & Beverly, B. L. (2015). Children with SLI exhibit delays resolving ambiguous reference. *Journal of Child Language, 42*(1), 180-195. doi:http://dx.doi.org.www2.lib.ku.edu/10.1017/S0305000914000038

Eyer, J. A., Leonard, L. B., Bedore, L. M., McGregor, K. K., Anderson, B., & Viescas, R. (2002). Fast mapping of verbs by children with specific language impairment. *Clinical Linguistics & Phonetics, 16*(1), 59-77. doi:http://dx.doi.org.www2.lib.ku.edu/10.1080/02699200110102269

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13,* 65-77.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation-2.* Circles Pines, MN: American Guidance Service.

Gray, S. (2003). Word-learning by preschoolers with specific language impairment: What predicts success? *Journal of Speech, Language, and Hearing Research, 46*(1), 56-67. doi:http://dx.doi.org.www2.lib.ku.edu/10.1044/1092-4388(2003/005)

Gray, S. (2004). Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research, 47*(5), 1117-1132. doi:http://dx.doi.org.www2.lib.ku.edu/10.1044/1092-4388(2004/083)

Gray, S. (2006). The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 49*(5), 955-969. doi:http://dx.doi.org.www2.lib.ku.edu/10.1044/1092-4388(2006/069)

Gray, S., & Brinkley, S. (2011). Fast mapping and word learning by preschoolers with specific language impairment in a supported learning context: Effect of encoding cues, phonotactic probability, and object familiarity. *Journal of Speech, Language, and*

Hearing Research, 54(3), 870-884. doi:http://dx.doi.org.www2.lib.ku.edu/10.1044/1092-4388(2010/09-0285)

Gray, S., Brinkley, S., & Svetina, D. (2012). Word learning by preschoolers with SLI: Effect of phonotactic probability and object familiarity. *Journal of Speech, Language, and Hearing Research, 55*(5), 1289-1300. doi:10.1044/1092-4388(2012/11-0095)

Gray, S., Pittman, A., & Weinhold, J. (2014). Effect of phonotactic probability and neighborhood density on word-learning configuration by preschoolers with typical development and specific language impairment. *Journal of Speech, Language, and Hearing Research, 57*(3), 1011-1025. doi:http://dx.doi.org.www2.lib.ku.edu/10.1044/2014_JSLHR-L-12-0282

Hansson, K., Forsberg, J., Löfqvist, A., Mäki-Torkko, E., & Sahlén, B. (2004). Working memory and novel word learning in children with hearing impairment and children with specific language impairment. *International Journal of Language & Communication Disorders, 39*(3), 401-422. doi:http://dx.doi.org.www2.lib.ku.edu/10.1080/13682820410001669887

Horohov, J. E., & Oetting, J. B. (2004). Effects of input manipulations on the word learning abilities of children with and without specific language impairment. *Applied Psycholinguistics, 25*(1), 43-65. doi:http://dx.doi.org.www2.lib.ku.edu/10.1017/S0142716404001031

Jackson, E., Leitao, S., & Claessen, M. (2015). The relationship between phonological short‐term memory, receptive vocabulary, and fast mapping in children with specific language impairment. *International Journal of Language & Communication Disorders*, doi:http://dx.doi.org.www2.lib.ku.edu/10.1111/1460-6984.12185

Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergarteners. *Language, Speech, and Hearing Services in Schools, 36*(1), 17-32.

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 53*(3), 739-756. doi:http://dx.doi.org.www2.lib.ku.edu/10.1044/1092-4388(2009/08-0248)

Kiernan, B., & Gray, S. (1998). Word learning in a supported-learning context by preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*(1), 161-171.

Leonard, L. B. (2014). *Children with specific language impairment.* Cambridge, MA: MIT Press.

McGregor, K. K., Licandro, U., Arenas, R., Eden, N., Stiles, D., Bean, A., & Walker, E. (2013). Why words are hard for adults with developmental language impairments. *Journal of Speech, Language, and Hearing Research, 41*(1), 161-171.

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *Int J Lang Commun Disord, 48*(3), 307-319. doi:10.1111/1460-6984.12008

McKean, C., Letts, C., & Howard, D. (2013). Developmental change is key to understanding primary language impairment: The case of phonotactic probability and nonword repetition. *Journal of Speech, Language, and Hearing Research, 56*(5), 1579-1594. doi:http://dx.doi.org.www2.lib.ku.edu/10.1044/1092-4388(2013/12-0066)

Oetting, J. B., Rice, M. L., & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech & Hearing Research, 38*(2), 434-445.

Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds Intellectual Assessment Scales and the Reynolds Intellectual Screening Test.* Lutz, FL: Psychological Assessment Resources, Inc.

Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech & Hearing Disorders, 55*(1), 33-42.

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research, 37*, 106-122.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals – Fourth Edition.* San Antonio, TX: The Psychological Corporation.

Seymour, H. N., Roeper, T. W., de Villers, J., & de Villers, P. A. (2005). *Diagnostic Evaluation of Language Variation – Norm-Referenced.* San Antonio, TX: The Psychological Corporation.

Storkel, H. L. (2004a). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics, 25,* 201-221.

Storkel, H. L. (2004b). Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research, 47*, 1454-1468.

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research, 49*, 1175-1192.

Storkel, H. L. & Hoover, J. R. (2010). An on-line calculator to compute phonotactic probability and neighborhood density based on child corpora of spoken American English. *Behavior Research Methods, 42(2), 497-506.* Available from PubMed Central.

Storkel, H. L., & Lee, S. (2011). The independent effects of phonotactic probability and neighborhood density on lexical acquisition by preschool children. *Language and Cognitive Processes, 26,* 191-211.

Storkel, H. L., Voelmle, K., Fierro, V., Flake, K., Fleming, K. K., & Swinburne Romine, R. (2017). Interactive book reading to accelerate word learning by kindergarten children with specific language impairment: Identifying an adequate intensity and variation in treatment response. *Language, Speech, and Hearing Services in Schools, 48,* 16-30. doi: 10.1044/2017_LSHSS-16-0014

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech & Hearing Research, 40*(6), 1245-1260.

Torrens, V., & Yagüe, E. (2016). The role of phonological working memory in children with SLI. *Language Acquisition.* doi:10.1080/10489223.2016.1187617

Wagner, R. K., Torgesen, J., Rashotte, C., & Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing* (2nd ed).

Windfuhr, K. L., Faragher, B., & Conti-Ramsden, G. (2002). Lexical learning skills in young children with specific language impairment. *International Journal of Language & Communication Disorders, 37*(4), 415-432. doi:http://dx.doi.org.www2.lib.ku.edu/10.1080/1368282021000007758

Appendix A

Outcome Data across Conditions

The purpose of Study 2 was to determine the optimal way to achieve 36 elaborated exposures to the study vocabulary words. Although this was not the focus of the current investigation, outcome data for children in each condition were compared descriptively. Figure 4 below shows children's outcomes in each of the three conditions. For treated vocabulary words, children in the maximize dose frequency condition (*n* = 4) learned 6 words on average (range = 1-13, SD = 5.60). Likewise, children in the balanced condition (*n* = 6) learned 6.7 words on average (range = 0-16, SD = 5.05). In contrast, children in the maximize dose condition (*n* = 4) learned 1.5 words on average (range = 0-2, SD = 1.00), which is much lower than children in the other two conditions. These results were compared to the number of untreated control words that children learned following treatment (i.e., words from set B for children learning set A words, and words from set A for children learning set B words). Children in all conditions correctly defined few (i.e., less than 3) untreated control words. However, error bars on the graph show that the standard error of measurement ranges for treatment vs untreated control words defined correctly for children in the maximize dose frequency and balanced conditions do not overlap, while ranges for the treatment vs untreated control words in the maximize dose condition do overlap. This suggests that appreciable learning is not occurring, since the treatment seems to have little to no additional effect on the number of words children learn.

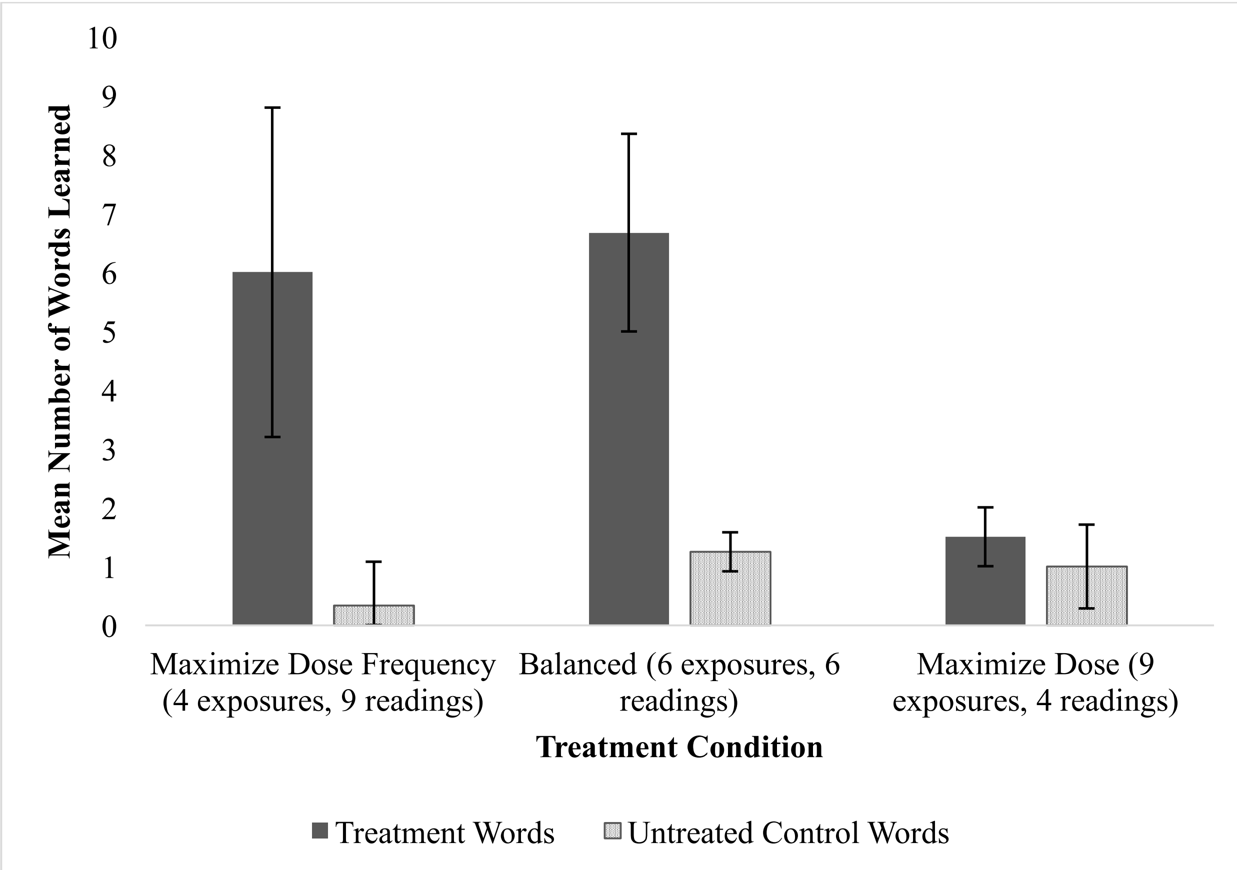

Figure 4. Differences in treatment response across study conditions for treatment and untreated control words. Error bars on the graph represent the standard error of measurement.

To further investigate the effectiveness of each treatment condition, three separate paired samples t-tests were conducted comparing the number of treatment and untreated control words correctly defined by children in each of the three conditions. The comparison for children in the balanced condition was significant, with fewer untreated control words (M = 0.33, SD = 1.00) than treatment words (M = 6.67, SD = 5.05) learned, $t(8) = -4.30$, $p < .05$. The comparison between untreated control words (M = 1.25, SD = 1.50) and treatment words (M = 6.00, SD = 5.60) learned was not significant, $t(3) = -1.9$, $p = 0.38$. For the maximize dose condition, the comparison between untreated control words (M = 1.00, SD = 1.41) and treatment words (M = 1.50, SD = 0.50) learned was not significant, $t(3) = -0.78$, $p = 0.50$. These non-significant findings may have been caused by the low statistical power resulting from having only 4 children in each condition. Descriptively, the *t* value is higher and the *p* value is lower for the children in the maximize dose frequency condition than for children in the maximize dose condition, indicating a more appreciable difference in learning.

Since children seemed to show appreciable learning in the maximize dose frequency and balanced conditions, these children's data were used in the current analysis. Children in the maximize dose condition were excluded from these analyses since their learning did not sufficiently exceed learning for untreated control words. This decision was made due to concerns that results would be confounded by some participants receiving a less-than-adequate treatment schedule. Importantly, children included in the analyses still showed a range of outcomes. Not all children who showed little learning were excluded, only those in the potentially less optimal treatment condition.

Appendix B

Study Word Characteristics

Word Set	Book	Word	Part of Speech	Word Length in Phones[1, 2]	Positional Average[1, 3]	Biphone Average[1, 4]	Number of Neighbors[1]
A	Harry and the Terrible Whatzit	damp	adjective	4	0.056	0.007	6
		discovered	verb	7	0.049	0.005	0
		furnace	noun	5	0.052	0.004	0
		gloomy	adjective	4	0.031	0.002	1
		swat	verb	4	0.064	0.003	7
		swung	verb	4	0.043	0.002	2
	Imogene's Antlers	advice	noun	5	0.019	0.000	0
		glared	verb	4	0.035	0.003	1
		overjoyed	adjective	5	0.006	0.001	0
		prodded	verb	4	0.057	0.004	2
		rare	adjective	3	0.068	0.008	16
		wandered	verb	5	0.070	0.008	2
	Otis	hauled	verb	3	0.029	0.001	5
		hooves	noun	4	0.029	0.000	0
		ripe	adjective	3	0.039	0.002	11
		sidelines	noun	6	0.060	0.002	0
		silky	adjective	4	0.088	0.008	5
		spotless	adjective	7	0.053	0.003	0
	Possum and the Peeper	clamor	noun	5	0.053	0.004	1
		grumbling	verb	6	0.033	0.005	1
		marsh	noun	4	0.053	0.010	5
		peering	verb	3	0.084	0.006	20
		racket	noun	5	0.068	0.006	1
		squinting	verb	6	0.057	0.005	0
	Shy Charles	embarrassed	adjective	7	0.028	0.002	0
		murmured	verb	4	0.048	0.001	0
		nervous	adjective	5	0.032	0.002	1
		scarlet	noun	7	0.051	0.004	0
		success	noun	6	0.055	0.002	0
		trembled	verb	6	0.038	0.004	0
B	Book! Book! Book!	gathered	verb	4	0.048	0.002	1
		heaved	verb	3	0.040	0.002	10
		pouted	verb	3	0.050	0.001	11
		ruffle	verb	4	0.038	0.002	3
		squawked	verb	5	0.038	0.004	1
		whinnied	verb	5	0.074	0.007	2
	Swimmy	gulp	noun	4	0.054	0.002	1
		invisible	adjective	8	0.022	0.002	0
		marvel	noun	5	0.049	0.007	2
		midday	noun	4	0.053	0.005	2
		swaying	verb	3	0.055	0.003	5
		swift	adjective	5	0.058	0.003	1
	The Bear Under the Stairs	awful	adjective	3	0.005	0.001	2
		crept	verb	5	0.065	0.005	1
		decided	verb	5	0.054	0.006	1

Word Set	Book	Word	Part of Speech	Word Length in Phones[1, 2]	Positional Average[1, 3]	Biphone Average[1, 4]	Number of Neighbors[1]
		haddock	noun	5	0.065	0.005	1
		noticed	verb	5	0.045	0.004	0
		tight	adjective	3	0.055	0.004	19
	The Caterpillar that Roared	gaze	verb	3	0.034	0.001	5
		horrified	adjective	6	0.040	0.002	0
		ripples	noun	4	0.050	0.005	4
		snuggled	verb	5	0.042	0.002	1
		surface	noun	5	0.051	0.002	1
		twitch	verb	4	0.029	0.004	4
	What Do You Do With a Kangaroo	flashing	adjective	4	0.034	0.005	4
		frayed	adjective	3	0.053	0.010	8
		smooth	adjective	4	0.037	0.001	0
		stale	adjective	4	0.052	0.009	11
		tailor	noun	4	0.063	0.004	4
		worn	adjective	4	0.052	0.005	2
Summary for Set A words:			**noun** = 9 **verb** = 12 **adjective** = 9	***M*** = 4.83 ***SD*** = 1.26 **Range** = 3-7	***M*** = 0.048 ***SD*** = 0.018 **Range** = 0.006-0.088	***M*** = 0.004 ***SD*** = 0.003 **Range** = 0.000-0.010	***M*** = 2.90 ***SD*** = 4.92 **Range** = 0-20
Summary for Set B words:			**noun** = 7 **verb** = 12 **adjective** = 10	***M*** = 4.3000 ***SD*** = 1.0876 **Range** = 3-8	***M*** = 0.047 ***SD*** = 0.014 **Range** = 0.005-0.074	***M*** = 0.004 ***SD*** = 0.002 **Range** = 0.001-0.010	***M*** = 3.57 ***SD*** = 4.30 **Range** = 0-19
Statistics comparing part of speech (chi square test for independence):		$X^2 =$ $p =$	2.00 0.20	-- --	-- --	-- --	-- --
Statistics comparing phonotactic probability and neighborhood density (independent samples t-test):		$t =$ $p =$	-- --	1.75 0.09	0.32 0.75	-0.06 0.95	-0.56 0.58

[1]Calculated from Storkel and Hoover (2010) using the child database. [2]Word forms were modified when the exact match was not found in the database (e.g., “hauled” was entered as “haul”). [3]Sum of the positional segment frequency of each sound (Storkel, 2004b) divided by the number of sounds. [4]Sum of the biphone frequency of each adjacent pair of sounds (Storkel, 2004b) divided by the number of adjacent sound pairs.